NBER WORKING PAPER SERIES

FORECLOSURE SPILLOVERS WITHIN BROAD NEIGHBORHOODS

Weiran Huang
Ashlyn Nelson
Stephen Ross

ABSTRACT

Most evidence on foreclosure spillovers identifies localized effects that are modest in magnitude, but these effects could multiply to larger aggregate effects across broad neighborhoods. We test this proposition developing a proxy for the fraction of mortgages in negative equity during the foreclosure crisis and estimating a difference-in-differences model for foreclosure. This proxy exploits the timing of foreclosures in each tract, and this within tract variation is not predicted by mortgage attributes, housing attributes or sales prices. Our estimates suggest that 61 percent of the increase in across tract dispersion in foreclosure filings can be explained by these spillover effects.

Weiran Huang
Department of Finance
City of New York
New York, NY
weiran.huang@gmail.com

Ashlyn Nelson
School of Public and Environmental Affairs
Indiana University
1315 E. Tenth St., Office 410D
Bloomington, IN 47405
ashlyn@indiana.edu

Stephen Ross
Department of Economics
University of Connecticut
341 Mansfield Road, Unit 1063
Storrs, CT 06269-1063
and NBER
stephen.l.ross@uconn.edu

Foreclosure Spillovers within broad Neighborhoods

A growing literature examines the spillover effects of foreclosures on local housing prices and the likelihood of future foreclosures.[1] Many studies have documented high contemporaneous rates of foreclosure and heavily depressed housing prices in neighborhoods that have experienced a large number of foreclosures including Immergluck and Smith (2006), Leonard and Murdoch (2009) and Rogers and Winter (2009) for housing prices and Goodstein et al. (2017) and Bradley et al. (2015) for foreclosures. Unfortunately, it is difficult to disentangle the effects of foreclosure rates at the neighborhood level from the effect of economic factors and household unobservables that contribute to those foreclosure rates.[2]

In response, many recent studies exploit high frequency data over space and time to identify the causal effect of foreclosure over very localized housing markets. These studies include high dimensional fixed effects to control for neighborhood unobservables and so identify the effect of foreclosures on housing sale prices located within a specified distance of the foreclosure, typically between 250 feet and one half mile (Schuetz et al. 2008; Harding et al. 2009; Campbell et al. 2011; Anenberg and Kung 2014; Hartley 2014; Gerardi et al. 2015).[3] These studies typically find modest, negative effects of nearby foreclosures on housing prices.[4]

---

[1] Other studies on neighborhood or peer spillovers include studies of crime (Glaeser, Sacerdote and Scheinkman 1996), employment (Bayer, Ross, and Topa 2008), welfare usage (Bertrand, Luttmer, and Mullainathan 2000), pre-natal care (Aizer and Currie 2004), health behaviors (Fletcher and Ross 2018) and academic outcomes (Fletcher, Ross and Zhang 2013). See recent surveys by Ioannides and Topa (2010) and Ross (2011).

[2] A related literature examines the concentration of high cost lending and foreclosure in disadvantaged and minority neighborhoods. For recent examples see Ghent, Hernández-Murillo, and Owyang (2014), Chan, Gedal, Been and Haughwout (2013), Chan, Haughwout and Tracy (2015), and Bayer, Ferreira and Ross (2016, 2017).

[3] Bayer, Ross and Topa (2008) argue that thin housing markets and limited information on micro neighborhoods make it difficult for households to systematically sort into specific houses or residential blocks conditional on broader neighborhoods, and support this assertion empirically with balancing tests.

[4] One exception is Lin et al. (2009) who find large housing price effects, but this difference may be because they do not control for neighborhood trends unlike many of the other studies cited.

A smaller literature investigates whether social networks or other non-price spillovers influence foreclosure rates within neighborhoods. Guiso et al. (2013) find that individuals who know someone who defaulted on their mortgage are 82 percent more likely to state their intention to default. Following the literature on housing prices, Towe and Lawley (2013) identify the causal effect on foreclosure of earlier foreclosures among an individual's 13 to 25 closest neighbors after conditioning on the broader neighborhood. Munroe and Wilse-Samson (2013) and Gupta (2019) both use instrumental variable approaches, random assignment of chancery-court judges to foreclosure cases and the contractual terms of adjustable-rate mortgages respectively, to identify the causal effect of foreclosure over and above any price effects. Both studies again focus on very localized effects, i.e. foreclosures within 0.1 of a mile of the household. All three studies find substantial evidence that earlier foreclosures raise the likelihood that other nearby households experience foreclosure. If foreclosure activity is sufficiently dense in a neighborhood, these localized network effects could spill over space and multiply, leading to higher average foreclosure rates for the surrounding neighborhood (Mian, Sufi and Trebbi 2015; Guren and McQuade 2020).

However, in spite of the growing evidence of causal effects of foreclosures on very localized housing markets, we do not know whether these localized spillovers and/or other neighborhood spillovers on housing price or neighborhood quality contribute substantially to the large conditional correlation between neighborhood foreclosure rates and the likelihood of future foreclosures in the same neighborhood. In one exception to this limitation, Gupta (2019) finds that foreclosures are explained by regional foreclosure rates using a share of mortgages that are both near negative equity and have resetting adjustable rates as an instrument. By exploiting resetting rates, the treatment effects identified are driven on the margin by changes in the subprime market across locations. As an alternative strategy that may capture the patterns of foreclosure spillovers

in the broader market, we exploit the high correlation between negative equity and foreclosure by developing an exogenous proxy for the share of households in negative equity for each neighborhood and quarter during the financial crisis. We then use predicted level of negative equity to estimate an aggregate effect of expected neighborhood foreclosure activity on the likelihood of future foreclosure notices. Therefore, rather than obtaining identification using detailed spatial variation at a given point in time, the identifying variation in our paper arises from variation over time within location exploiting the increases in negative equity that occurred as housing prices fell during the crisis.[5]

Specifically, we develop a prediction for the fraction of neighborhood housing units purchased in the run up to the housing crisis that are in negative equity during a given year and quarter. This prediction is based on a metropolitan housing price index and so does not depend upon neighborhood housing prices. The prediction also does not depend upon the Combined Loan to Value (LTV) ratios selected by the buyers in that neighborhood or during time period of the purchase because the LTV distribution for a given neighborhood omits transactions within that neighborhood and is based on the entire pre-crisis period.[6] This prediction exploits the timing of when individuals purchased housing in each neighborhood and combines this information with the time path of housing prices and the distribution of LTV's for the overall housing market.[7] For example, if a neighborhood happens to have more transactions near the peak of the market, then this neighborhood will be predicted to have more housing units in negative equity early in the

---

[5] The models that exploit spatial variation usually include location by time period controls so that they can identify the proximate effect of the occurrence of new foreclosure, while our paper will include location fixed effects and exploit expected increases in the incidence of negative equity during the crisis within each location.

[6] The combined loan to value ratio is the ratio of the total value of all liens on the property at the time of purchase to the sales price of the property. For convenience, LTV is used as a short-hand for this ratio.

[7] This identification strategy is similar to Anenberg and Kung (2018) and Guren (2018) who instrument for house listing price using MSA level appreciation between when the house was purchased and the listing date comparing people who are listing their house at the same time, but purchased the house at different times.

crisis. We also verify that there is no within neighborhood correlation between the volume of housing transactions in a given quarter and the attributes of the mortgages originated, homes purchased and even the sales prices during that quarter. While borrower/buyer attributes like race and income are unobserved, sales price and housing attributes like square feet are usually correlated with income and loan to value ratios are typically higher for African American borrowers.

We specify a model of foreclosure as a function of the predicted share of mortgages experiencing negative equity within a census tract controlling for contemporaneous LTV for the individual mortgage plus purchase quarter and year by census tract fixed effects and purchase quarter and year by crisis quarter and year fixed effects. This fixed effect structure allows each time cohort of home purchase loans in the market to have a unique time path of foreclosures during the crisis and allows for heterogeneity in the unobservables of those buyers over purchase quarter within each census tract in terms of average foreclosure risk. In terms of identification, the predicted negative equity instrument varies across crisis quarters as housing prices fall and varies across census tracts because tracts differ in their initial equity position based on the timing of purchases leading up to the crisis. The model can then be interpreted as a difference-in-differences analysis because the fixed effects absorb tract and crisis quarter variation, and the instrument is identified by the interaction of the tract and crisis quarter variation. The key functional form or non-linearity that our instrument exploits is the negative equity threshold, which is presumed to have a discontinuously large effect on the likelihood of foreclosure relative to other changes in the level of home equity in a neighborhood.

Specifically, we estimate models for whether a housing unit received the first foreclosure notice, or Notice of Trustee Sale (NOT), in a given quarter using a sample of single-family housing

units that sold as arm's length transactions between the third quarter of 2001 and the second quarter of 2006 in San Diego County. Unlike judicial foreclosure states, the administrative process of foreclosure in California is typically initiated by lenders issuing a NOT when the borrower is approximately 120 days behind on their mortgage payments. The first foreclosure notice variable is based on foreclosure notices on those units issued between the third quarter of 2006 and the second quarter of 2010.[8] The likelihood of receiving a first foreclosure notice is then allowed to depend upon the stock of foreclosures since the on-set of the foreclosure crisis. Our instrument for the stock of foreclosures is strongly predictive of the total number of units in a census tract that experienced their first NOT by the end of the quarter. We estimate both reduced form models and Instrumental Variable (IV) models. The later models require the assumption that neighborhood levels of negative equity only influence the likelihood of a first NOT through the number of units in foreclosure, i.e. units having already received a NOT since the on-set of the crisis, after controlling for the actual negative equity level of each mortgage.

The estimated spillover effects are quite large. A one standard deviation change in the share of units in negative equity is associated with an 85% increase in the likelihood of a unit purchased during our pre-crisis period entering foreclosure, receiving a first NOT, in the next quarter over a base of 0.7 percentage points using the reduced form estimates. Similarly, a one standard deviation increase in the number of units having received a NOT implies a 58% increase in the likelihood of a mortgage in the tract entering foreclosure in a quarter. The magnitude of both the reduced form

---

[8] While our model relates current foreclosure likelihood to a proxy for the anticipated likelihood of foreclosure levels at that time in the neighborhood, our model does not explicitly solve the reflection problem where Manski (1993) argues that it is impossible to disentangle the effects of exogenous neighborhood or peer attributes from endogenous effects arising from the choices made by those neighborhood peers. We face the same limitation in that we cannot explicitly distinguish between the effects of neighborhood negative equity levels and the effects of preceding foreclosure decisions by residents in the neighborhood. However, we anticipate that both neighborhood negative equity levels and neighborhood foreclosures both operate in the same direction, and our proxy captures the combined effect of both spillover mechanisms. We obtain lower bounds on the overall effects of neighborhood negative equity on foreclosure by estimating reduced form models.

and IV estimates are very stable as we add controls for mortgage attributes. The magnitude of the IV estimates are also unaffected by the addition of controls intended to capture changes in the economic circumstances of each census tract over the crisis.[9] The fact that the magnitude of our estimates are unaffected by these controls provides further evidence that our results are not driven by unobservables at the census tract level that have heterogeneous effects over time during the crisis. We observe very similar IV estimates for both smaller (census block group) and larger (zip code) geography.

We also assess the magnitude of these effects by examining the increase in across neighborhood dispersion in the stock of mortgages having received foreclosure notices. Using our IV estimates, we find that the increase in the dispersion of ever received a foreclosure notice across neighborhoods can explain 61% of the increase in the dispersion of new foreclosure filings during the crisis, consistent with neighborhoods following substantially different foreclosure paths in part because past foreclosures are having spillover effects on other mortgages. Further, even as housing prices recover, these neighborhood spillover effects contribute to persistently high rates of new foreclosures due to the large stock of past foreclosure notices.

We also examine whether the spillover effects vary by neighborhood type. We interact the predicted fraction of units in negative equity with the tract share black, share Hispanic or share of households in poverty either including those interactions directly in the reduced form estimates or using them as instruments for the interaction of neighborhood foreclosure levels with the same tract attributes. The reduced form estimates suggest that the foreclosure effects of negative equity are much larger in these disadvantaged neighborhoods with the differences primarily being driven

---

[9] First, we allow for non-parametric trends in foreclosure over tract attributes that were observed prior to the crisis interacting crisis year dummy variables with pre-determined census tract observables. Then, we develop a Bartik style index based on the industry employment shares of tract residents and the employment shocks experienced in the State of California during the crisis.

by neighborhood share Hispanic. However, the IV estimates are relatively constant across different types of neighborhoods. These results are consistent with negative equity being a stronger predictor of foreclosure in disadvantaged neighborhoods so that the larger reduced form estimates are deflated by a larger first stage, but that the actual spillover effects are relatively similar across neighborhoods.

In summary, we find large spillover effects of foreclosures in the San Diego housing market during the housing crisis with increases in negative equity and overall foreclosure rates being associated with large increases in the baseline rate of new foreclosure notices. The magnitude of these results are very stable to the inclusion of observed mortgage risk factors, robust to the inclusion of controls intended to capture census tract trends during the crisis and robust to alternative neighborhood geographies. The results are also broad based occurring in both disadvantaged neighborhoods and neighborhoods with smaller fractions of minority residents and families in poverty.

Methodology

Predicting Neighborhood Levels of Negative Equity

In this section, we construct our instrument or proxy for neighborhood foreclosure risk $\widehat{NE}_{gt}$ as the predicted level of negative equity in each neighborhood g and in crisis quarter and year t. Specifically, we calculate the fraction of home purchases likely to be in negative equity in each current quarter at the census tract level based on the quarter of purchase, the market wide distribution of initial Combined Loan to Value (LTV) ratios, and time pattern of market wide housing prices.

First, we develop the market wide housing price index by estimating a traditional hedonic sales price model:

$$P_{jp} = \beta W_j + \delta_p + \varepsilon_{jp} \tag{1}$$

where P_{jp} is the logarithm of the sales price of house j at purchase quarter and year p, W_j is a vector of housing attributes of house j, and δ_p is a purchase quarter and year fixed effect. The resulting price index $(\bar{P})$ is:

$$\bar{P}_p = \frac{Exp(\bar{\delta}_p^5)}{Exp(\bar{\delta}_1^5)} \tag{2}$$

where $\bar{\delta}_p^5$ is a five quarter moving average of estimated quarter fixed effects $\hat{\delta}_p$ centered on quarter p. This price index is used to calculate the current LTV of individual housing units based only on housing price movements ignoring amortization, which endogenously depends upon mortgage product choice and payment history. We use market wide price indices even for the current LTV on individual units since using a neighborhood level price index would capture the effects of foreclosures on local housing prices, which is a phenomenon that we intend to include within our assessment of the magnitude of broad neighborhood spillovers.

Second, we define and calculate the market wide distribution of initial LTV ratios using relatively small LTV bins (b) for computational convenience. This distribution is calculated neighborhood by neighborhood in order to omit the transactions actually occurring in neighborhood g, as follows:

$$D_{b,-g} = Fr\left(\overline{LTV}_{b-1} < LTV_j \leq \overline{LTV}_b | j \notin \Omega_g\right) \tag{3}$$

where Fr represents the fraction of individual property sales (j) that fall between the lower bound $\overline{LTV}_{b-1}$ and the upper bound $\overline{LTV}_b$ of each LTV bin b, and Ω_g represents the set of transaction in neighborhood g. For each LTV bin, we calculate whether the loans in that bin would be in negative equity in crisis quarter and year t if the loans had been originated in pre-crisis quarter p. In practice,

estimates are very similar whether or not own tract is omitted and whether the LTV distribution is held constant over time or allowed to vary by quarter.

Specifically, using the price index from equation (2), a mortgage in a given bin is predicted to be in negative equity (ignoring amortization) if

$$\overline{LTV}_{b-1} * \frac{\bar{P}_p}{\bar{P}_t} > 1 \tag{4}$$

Then, for each quarter during the crisis (t), we calculate a predicted fraction of home purchases during the run up to the crisis that are likely to be in negative equity. We create this fraction by summing over all LTV bins the product of the share of transactions (D_{bp}) associated with a specific purchase quarter LTV bin (b) and a binary indicator (I) for whether the predicted, current LTV in equation (4) is in negative equity over all bins for a given purchase quarter/crisis quarter combination. Next, we multiply this sum (the fraction of purchases predicted to be in negative equity) by the number of transactions N_{gp} in that purchase quarter and neighborhood to obtain the predicted number of houses in negative equity in each neighborhood (g), purchase quarter (p), and current quarter (t). We then sum these totals over all purchase quarters and scale by the total number of transactions in the neighborhood during the entire pre-crisis period (N_g):

$$\widehat{NE}_{gt} = \frac{1}{N_g} \sum_{p=1}^{P} N_{gp} \left(\sum_{b=1}^{B} D_{b,-g} I \left(\overline{LTV}_{b-1} * \frac{\bar{P}_p}{\bar{P}_t} > 1 \right) \right) \tag{5}$$

where P is the total number of quarters in the pre-crisis period and B is the number of bins.

Other than the omission of the transactions from the unit's own neighborhood, the term in parentheses only depends upon the purchase quarter and the current quarter because both the price index and LTV distribution are based on the entire housing market, and all results are robust to estimating the model without omitting own neighborhood transactions. Therefore, again ignoring omission of own tract, one can write $\widehat{NE}_{gt}$ as a non-linear function $NE\left(\underline{N}_g, \bar{P}_t\right)$ where $\bar{P}_t$ is the

price level in the relevant crisis quarter and $\underline{N}_g$ is a vector of transaction volumes N_{gp} over all purchase quarters. Therefore, the across tract variance in $\widehat{NE}_{gt}$ arises entirely from the timing of purchases in each neighborhood and our exploitation of discrete effects associated with the negative equity threshold.[10]

Estimation Equations

To examine whether individual housing units are more likely to be foreclosed in neighborhoods with higher foreclosure rates, we regress an indicator for receiving a first Notice of Trustee Sale (NOT) on our predicted level of negative equity, $\widehat{NE}_{gt}$ or $NE(\underline{N}_g, \bar{P}_t)$. Our model includes neighborhood g by purchase quarter p (during the pre-crisis period) fixed effects (α_{gp}) to allow for differences in foreclosure rates across each quarter cohort of mortgages within each neighborhood. The model also include crisis quarter t by purchase quarter p fixed effects (δ_{pt}) so that each purchase quarter cohort of loans has a different market wide time profile of foreclosure over the crisis period, and the current loan to value ratio of the property.

$$NOT_{igpt} = \beta NE(\underline{N}_g, \bar{P}_t) + \theta LTV_{igpt} + \alpha_{gp} + \delta_{pt} + \varepsilon_{igpt} \qquad (6)$$

where NOT_{igpt} is whether the housing unit i received a first NOT in crisis quarter t and LTV_{igpt} is vector including controls for the current combined loan to value ratio for the mortgage itself capturing the current level of negative equity, again ignoring amortization.

The resulting model resembles a traditional differences-in-differences model in a panel data context such as $y_{it} = Z_t X_i + \alpha_i + \delta_t + \varepsilon_{it}$ where Z_t is a time varying attribute, X_i contains

---

[10] Note that our instrument employs $\overline{LTV}_{b-1}$, the lower bound of the LTV bin, in order to obtain the most conservative estimate of negative equity. Similar results arise using the upper bound of the bin to calculate predicted LTV. We also have calculated the predicted share of mortgages at higher levels of negative equity, as well as the predicted average amount of negative equity. These instruments also have predictive power, but the share negative equity is the strongest instrument and results are similar after the inclusion of these additional instruments. Therefore, we focus on the simpler model with one instrument for stock of units having experienced foreclosure.

cross-sectional variation, and the model is identified by the interaction of cross-sectional/individual information with aggregate time varying information conditional on both individual and time fixed effects. In our case, the vector of tract transaction volume over time represents the cross sectional information about each tract, the price level for the current quarter is the time varying information, and *NE* is a non-linear function combining the two variables based on our specific assumption concerning the importance of negative equity in driving foreclosures. If N_{gp} is higher earlier in the pre-crisis period when $\bar{P}_p$ is lower, then more of the houses in that neighborhood will be in a better equity position at the start of the crisis, which will delay increases in the predicted fraction of housing units in negative equity until later in the crisis. In this way, identification of the model is driven by the timing of housing units purchase in the neighborhood in the run up to the crisis. The instrument does, however, rely on the proposed empirical relationship between foreclosure and whether households are in negative equity. If in reality entering negative equity does not have a discrete effect on borrower default above and beyond continuous changes in equity position, our instrument would be expected to have minimal power.

If we are willing to impose the assumption that neighborhood rates of negative equity only influence foreclosures through the number of on-going foreclosures, we can estimate an Instrumental Variables (IV) specification using Two-Stage Least Squares. The first stage regresses a variable for the number of mortgages in tract *g* that have received a first NOT by crisis quarter *t* ($\overline{NOT}_{gt}$) on the same controls as in equation (6)

$$\overline{NOT}_{gt} = \beta'\widehat{NE}_{gt} + \theta'LTV_{igpt} + \alpha'_{gp} + \delta'_{pt} + \varepsilon'_{igpt} \qquad (7)$$

The second stage equation simply replaces $\widehat{NE}_{gt}$ in equation (6) with the predicted number of first NOT's. While $\overline{NOT}_{gt}$ only varies at the *{g,t}* level, we estimate this in a sample of observations

for explaining first notice of trustee sale NOT_{igpt} in equation (8) so that β' measures the average strength of the first stage for the relevant second stage sample

$$NOT_{igpt} = \tilde{\beta}\widehat{NOT}_{gt} + \tilde{\theta}LTV_{igpt} + \tilde{\alpha}_{gp} + \tilde{\delta}_{pt} + \tilde{\varepsilon}_{igpt} \tag{8}$$

Standard errors are clustered at the neighborhood level for all models, which addresses concerns that key variables in equations (7) and (8) only vary at the *{g,t}* level.[11]

Data and Sample

We use DataQuick Information Systems Inc. provided data on all home purchase transactions and all Notice of Trustee Sale recordings from the first quarter of 2001 through second quarter of 2010 collected from the San Diego County assessor's office. All home purchase transactions are recorded capturing the sale price and date, as well as information on both the primary and up to two subordinate liens/mortgages securing the purchase; and an associated assessors file contains the address of the housing unit and property attributes. California is an administrative/non-judicial foreclosure state, and a Notice of Trustee Sale (NOT) informs homeowners that their homes may be sold at a California public trustee foreclosure auction at any time after 21 days following the date that the notice is recorded in the county public record.

Housing price indices are created by estimating a hedonic model with purchase quarter by year fixed effects for the entire sample period. The hedonic attributes include the age of the unit, lot size, the square footage of the housing unit, the number of bedrooms, the number of bathrooms, whether the unit is a condominium or coop within a larger structure, and whether the housing unit is a 2 to 4 family. Since our data ends in the second quarter of 2010, the 5 quarter moving average price index for the first and second quarter of 2010 are based on increasing the weight on the

---

[11] Alternatively, equation (7) can be estimated with a sample at the *{g,t}* level and weighted by the number of observations in each cluster for the sample in equation (8), and similar results are obtained. In Table 4, we also present the results for an unweighted model in a *{g,t}* level sample and the instrument has significant power

second quarter to 2 and 3, respectively, to replace the unavailable (at the time of that we purchased the transaction data) 3rd and 4th quarters of 2010.

For our sample of housing units/home purchase mortgages, we select every single-family housing unit transaction excluding condos between the third quarter of 2001 and the second quarter of 2006 since housing units that had been held for longer than 5 years prior to the crisis had substantially more time to build up housing equity. If a housing unit was sold twice or more during this pre-crisis period, the most recent housing transaction is retained in the sample, and the earlier transactions are dropped so that our observations follow the home purchase mortgage of the individual who owns the home at the onset of the crisis.[12] We also drop any mortgages that were issued a NOT prior to the third quarter of 2006 so that we can focus on the risk of receiving a first foreclosure notice during the crisis period when housing prices were falling and equity was eroding.[13] We also drop all non-Arm's length transactions because we do not have an accurate measure of the value for calculating a combined loan to value ratio. Our final sample of transactions contains 121,185 single-family housing units purchased during the pre-crisis period that did not receive a NOT during the pre-crisis period.

Finally, for the foreclosure regression sample, we create a panel with one observation per housing unit for every quarter between the third quarter of 2006 and the second quarter of 2010. A NOT dummy variable is created that is zero for every quarter starting in the third quarter of 2006 until the mortgage/housing unit has a first NOT recorded in the county assessor's office. The NOT variable is set to one in this first NOT quarter, and all further quarters are dropped from the sample

---

[12] While we observed mortgage refinance behavior, we do not consider changes in equity due to refinancing because those decisions are likely endogenous to the market circumstances that arose following the home purchase. We also do not observe the amortization terms on individual mortgages, and so simply use the initial combined loan to value ratio for calculating the negative equity on individual mortgages.

[13] This sample restriction eliminates only a small fraction of the sample and has no impact on the results.

for that housing unit. If a housing unit does not receive a NOT during the crisis period, all quarters are retained and the NOT variable is zero for all quarters. This data structure provides the linear probability equivalent to a simple proportional hazard model.[14]

Table 1 presents descriptive statistics for the regression sample of housing units by crisis quarter, even though some variables are fixed over quarters at their values at the time of the purchase transaction for each housing unit.[15] The first row is the dependent variable, which is 1 only when the first NOT is recorded in that quarter. The rate of first NOT is quite low at 0.72 percent since it captures the flow of new NOT recordings. The fraction NOT, which captures the share of pre-crisis transactions in a census tract that have received a first NOT between the onset of the crisis and the end of the current quarter, provides a better sense of the level of foreclosure during the crisis period. The mean of this variable over all mortgages and quarters is 4.2 percentage points with a standard deviation of 4.8 suggesting sizable foreclosure rates and substantial variation across tracts and over time in those rates. The average tract rates of negative equity are quite high at 40 percent, also with a high standard deviation across census tracts of 25 percent. The predicted values of tract negative equity moderately under predicts the levels of the actual negative equity rates due to our use of the lower bound of the LTV bins, but they are highly correlated with the actual rates.

The Table also shows the means for the current combined LTV dummy variables, the transaction and mortgage attributes and the 2000 census tract attributes used in the regression models. During the sample period, 14% of mortgages have negative equity levels between 10 and 30 percent (LTV's between 110 and 130 percent), while 15% have negative equity levels above

---

[14] Logit estimation using data formatted in this way is mathematically equivalent to the standard proportional hazard model.

[15] We measure foreclosure and negative equity at the census tract level using 2000 census tract definitions.

30 percent. Most of the transactions are sales of existing housing that was built more than two years ago, 64% of the principle mortgages are adjustable rate mortgages, and over 44% of the home purchase transactions include subordinate debt. These high rates of adjustable rate mortgages and subordinate debt are typical of California's high cost housing market, but less representative of the U.S. overall. Over 22 percent of mortgages had initial combined Loan to Value (LTV) ratios over 95, but less than 3 percent had initial LTV's over 100. Using the 2000 census data, the typical share black or share households in poverty in a census tract is around 5 percent, while the average percent Hispanics is 21 percent.

Evidence on Identification

The model above is identified based on the timing of transactions up to the second quarter of 2006. The natural concern with this source of variation is that changes in the types of mortgages being issued in some tracts that experience increases in transaction volume could be different than the changes in the types of mortgages being issued in tracts that have flat or decreasing transaction volume over time. Table 2 presents means for individual cohort years of loans separately for tracts that experienced low, medium and high increases in predicted negative equity, i.e. tracts with different patterns of transaction volume over time.[16] Column 1 shows the bottom tercile of tracts over increases in predicted negative equity and column 3 shows the top tercile. Each panel shows the tercile means by cohort for a specific mortgage or housing unit attribute. The share of adjustable rate mortgages is increasing across cohorts until the 05-06 cohort of mortgages (third quarter 05 through second quarter 06). The share of subordinate debt increases with time throughout the period, along with the share of loans with LTV's greater than 95, but less than 100. The number of units sold that were less than two years old falls off dramatically in 04-05. However,

---

[16] Our crisis period starts in the 3rd quarter of 2006 so years from the 3rd quarter of one year to the 2nd quarter of the next.

these trends all appear in both the bottom and the top terciles based on increases in tract negative equity during the crisis. Table 2 does not indicate any strong or systematic differences in the changes in the composition of loans over time between tracts that had large and small increases in the fraction of units in negative equity.

We provide a more formal test for this concern examining whether tracts exhibit short-run balance over transaction volume, which is the source of our identifying variation. Specifically, we create a dependent variable for tract-quarter relative transaction volume that is the number of arm's length transactions in a tract and quarter divided by the median number of quarterly arm's length transactions in the tract, and regress this variable upon the quarterly mean mortgage and transaction attributes for each tract: whether adjustable rate mortgage, whether the transaction includes subordinate debt, dummies for loan to value ratio thresholds and whether the sale involves new construction. Next, we estimate an extended model that also adds mean hedonic attributes: lot size, square feet of living space, number of bedroom and number of baths, to the balance test, and finally a model that adds the log of sales price. Unfortunately, we do not observe borrower/buyer attributes, but loan to value ratios are typically correlated with race and both the price and the size of the house is almost certainly strongly correlated with income. Unlike the models in equations (6) through (8), relative number of transactions only varies at the tract by purchase quarter and year level. So, the model cannot include tract by purchase quarter and year fixed effects, and instead includes tract by purchase year fixed effects and purchase quarter and year fixed effects.

These balancing test results using our key controls for mortgage attributes are shown in Table 3 column 1. In column 2, we present the balancing test model including both mortgage and hedonic attributes, and in column 3 also add the sales price variable. Consistent with balance over covariates, none of the individual coefficient estimates are statistically significant except for

number of bathrooms, which is only significant at the 10% level. One rejection at the 10% level with 14 covariates is expected based on type 1 error. The F-statistic for the mortgage attributes in column 1 is only 0.60 with a probability of incorrectly rejecting the null of 0.74, and in column 2 and 3 F-statistics associated can reject the null with only 0.50 and 0.40 probability of type 1 error respectively, all far from significance. Finally, as shown later, the magnitudes of our estimated effects are robust to the inclusion of the key mortgage attributes, which provides further evidence of balance to support identification. We focus on balancing tests rather than pre-trends because the foreclosure rates prior to the on-set of the crisis are far too low to provide any power for rejecting the assumption of parallel pre-trends, and one or two quarter leads of predicted negative equity in an event study framework are too highly correlated with predicted current negative equity to be informative.[17]

Finally, Table 4 shows that the instrument has power to explain both the actual levels of negative equity and the share of mortgages receiving initial foreclosure notices. The predicted negative equity variable has a coefficient estimate of 0.75 in the model for the actual share of units in negative equity consistent with a very strong relationship between predicted and actual levels of negative equity. We find a similarly strong and statistically significant relationship between predicted negative equity in a given quarter and the share of mortgages having received foreclosure notices or NOT's by the end of that quarter. A one standard deviation increase in predicted share in negative equity is associated with a 5.9 percentage point increase in rates of ever having received a foreclosure notice relative to a base rate of 4.2 percent in our sample.

---

[17] Current predicted negative equity always wins the horserace against the lead of predicted negative equity in predicting higher rates of first NOT, but the estimation results show clear signs of serious multi-collinearity and so should not be trusted.

Results

Table 5 presents the reduced form first NOT model estimates for the predicted neighborhood negative equity measure plus the estimates on the current combined LTV dummy variables. The effect of tract levels of predicted negative equity on the likelihood of receiving a first NOT in a given quarter is substantial. A one standard deviation increase in predicted negative equity levels is associated with a 0.61 percentage point increase in the likelihood of receiving a first NOT relative to an average likelihood of 0.72 percent, or 85 percent of the average incidence. The magnitude of the estimated effect is very robust to the inclusion of initial mortgage attributes, essentially unchanged, even though the estimates on those controls are strong predictors of receiving a first NOT and dramatically erode (by approximately 50%) the magnitude of the estimates on the current LTV of the actual mortgage.

Table 6 presents the IV estimates regressing first NOT on the endogenous fraction of units with an NOT in the census tract, as well as the first stage estimates from the regression of fraction NOT on predicted levels of negative equity. The effects of fraction of mortgages having received an NOT are also highly significant and sizable. A one standard deviation increase in fraction ever NOT is associated with a 0.42 percentage point increase in the likelihood of any mortgage receiving a first NOT, about a 58 percent increase over the sample incidence of 0.72 percent. As in Table 4, the instrument is a powerful predictor of fraction NOT with an F-statistic of 32.[18] In fact, the coefficient on predicted negative equity is 40 percent larger in the transaction by crisis quarter sample as compared to the estimate from the tract by crisis quarter sample in Table 4. This large effect likely arises because the transaction sample places more weight on the tracts with more

---

[18] Kleibergen-Paap rk Wald F statistic allows for heteroskedastic and clustered errors. However, no formal critical values have been calculated for weak instruments outside of the iid case, and so Baum, Schaffer and Stillman (2007) recommend using the standard threshold of 10.

transactions where our predicted measure of negative equity more accurately captures actual levels of negative equity and foreclosure, i.e. suffers from less measurement error. As with the reduced form model, the magnitudes of these IV estimates are very robust across models without and with the additional transaction and mortgage controls.

Explaining the Increase in across Tract Dispersion

We next assess whether these estimated spillover effects can explain a substantial fraction of the large dispersion in foreclosure rates across neighborhoods that arose during the housing crisis. Table 7 presents the across tract dispersion of key variables by crisis year. The standard deviation in the incidence of first NOT recordings rises from 0.0041 in the first year after the crisis to an annual maximum of 0.0110 between the 3rd quarter of 2008 and the 2nd quarter of 2009 (the year that housing prices reach bottom), and represents an increase in the dispersion of foreclosure of 0.0069 or a 168 percent increase. Turning first to the reduced form analysis, the standard deviation of the fraction of pre-crisis mortgages predicted to be in negative equity rises from 0.0145 to 0.0613 during the same period. Multiplying the change in the standard deviation by the coefficient estimate on fraction in negative equity 0.031 from the first column of Table 5 yields 0.0014 implying that changes in the dispersion of negative equity can explain 20% of the increase in the standard deviation of first NOT recordings.

Turning to the IV estimates, the increase in the standard deviation between 2006-07 and 2008-09 in fraction of units having ever received a NOT is 0.0484. The IV estimate of the effect on the incidence of first NOT is 0.087. Multiplying the standard deviation change by the effect estimate implies an increase in the standard deviation of first NOT of 0.0042, or the increase in the dispersion of the stock of past mortgages having received a NOT can explain 61% of the 0.0069 change in the dispersion of the incidence of first NOT recordings. The larger share explained by

the IV estimates is not surprising given that the increase in the dispersion of ever NOT rates is substantially larger than the increase in the dispersion of predicted negative equity rates, and these large increases in ever NOT dispersion is expected in a world with foreclosure spillovers.

Table 7 also shows that housing prices stabilized between 08-09 and 09-10 rising by 2 percent, and the incidence of new NOT recordings falls from 1.1 to 0.8 percent. However, the base level of total units having received foreclosure notices rises from 6.7 to 10.3 percentage points. Applying the estimated effect of the stock of foreclosures implies (the IV estimate) an increase in the incidence of new units receiving notices of approximately 0.3 percentage points. Therefore, while a naive comparison of the changes between 08-09 and 09-10 implies that the stabilization of housing prices led to only a modest change in the rate of new first foreclosure recordings from 1.1 to 0.8 percentage points, after accounting for neighborhood level spillovers the stabilization of housing prices may have led to 0.6 percentage point reduction in new recordings relative to the rate that would have occurred without stabilizing housing prices. Mortgage default decisions should rationally be based on expectations of future housing prices, and so stabilizing housing markets would be expected to have very large effects on foreclosure rates. Those expected reductions in new foreclosures may have been by delayed by spillovers effects from the existing stock of foreclosures.

Heterogeneity across Tracts

Panel 1 of Table 8 presents estimates for reduced form models where predicted negative equity is interacted with the tract share of residents who are black, who are Hispanic or the tract share of households in poverty. Panel 2 of Table 8 presents equivalent IV estimates where these interactions serve as instruments for interactions of fraction NOT with the same neighborhood attributes. The reduced form estimates suggest that the foreclosure spillover effects of negative

equity are much larger in these disadvantaged neighborhoods with the differences primarily being

driven by neighborhood share Hispanic. The average effect of negative equity from Table 5 is 0.62

percentage points relative to a base frequency of first NOT notices of 0.72. A one standard

deviation change in percent Hispanic increases the effect by 0.14 percentage points. However, the

IV estimates are relatively constant across different types of neighborhoods with small estimates

on the interactions and level estimates that are comparable in magnitude to the estimates from

Table 6. These results are consistent with negative equity levels being a stronger predictor of

overall mortgage foreclosure (fraction NOT) in disadvantaged neighborhoods (likely

disadvantaged households are more vulnerable to declines in equity) so that the larger reduced

form estimates are deflated by a larger first stage. As a result, even though homeowners in

disadvantaged neighborhoods may be more sensitive to negative equity levels, the estimated

spillover effects of foreclosure are relatively stable across neighborhood types.

Robustness to Controls for Tract Trends during Crisis

Next, we extend the model to allow the time path of foreclosure in each neighborhood

during the crisis to vary systematically with pre-determined neighborhood observables (Z_g), see

Table 1 for the list of neighborhood variables.

$$NOT_{igpt} = \beta \widehat{NE}_{gt} + \theta LTV_{igpt} + \omega_t Z_g + \alpha_{gp} + \delta_{pt} + \varepsilon_{igpt} \tag{9}$$

where ω_t is a vector of dummy variables associated with each year of the crisis period.

These estimates are shown in Table 9. The inclusion of tract trends substantially reduces

the reduced form estimates of the effects of predicted negative equity in panel 1, perhaps because

bias from measurement error in our proxy is exacerbated by these additional controls. However,

the IV estimates that address potential measurement error are very similar to the estimates without

controls for trends on observables. The IV estimates from Table 6 are 0.087 and 0.089 without and

with controls for mortgage attributes, while the trend model IV estimates in Table 8 panel 2 are 0.087 and 0.079 without and with controls. Allowing for crisis period trends associated with location information that was available prior to the crisis does not change the basic conclusions of our paper. However, we prefer the models without these trends given the robustness of the second stage estimates and the fact that the inclusion of trends weakens the power of our instrument leading to F-statistics around 8, just below the typical threshold of 10 to avoid concerns about weak instruments.

As a second robustness test, we develop a proxy for the economic health of each neighborhood during the crisis period in order to capture economic trends that may have not been predicted by tract observables. Specifically, we follow Brunner, Ross and Washington (2011) using the industry of employment for tract residents based on 2000 census data to develop a Bartik style index of the economic shocks faced by residents.[19] Specifically, we use tract g employment shares in each industry k (s_{gk}) to weight statewide percentage changes in California industry employment from the Bureau of Labor Statistics for the two quarters preceding each crisis quarter t (G_{tk}), or

$$B_{gt} = \sum_{k=1}^{K} s_{gk} G_{tk}$$

where $G_{tk} = (E_{tk} - E_{t-2k})/E_{t-2k}$, E_{tk} is the statewide employment in quarter t and industry k, and K is the total number of industries on which the shares are based. We then estimate a model that includes the Bartik index as a control and interacts that index with current LTV since we expect foreclosure in response to negative equity to be affected by negative economic shocks, see for example Bayer, Ferreira and Ross (2016).

---

[19] The industries identified at the tract level from the 2000 decennial census include agriculture and mining, construction, manufacturing, wholesale trade, retail trade, transportation, information, finance, professional services, education services, art and recreation services, other services and public administration. The agriculture and mining category is omitted from our calculations.

$$NOT_{igpt} = \beta \widehat{NE}_{gt} + \theta_1 LTV_{igpt} + \theta_2 B_{gt} + \theta_3 B_{gt} * LTV_{igpt} + \alpha_{gp} + \delta_{pt} + \varepsilon_{igpt} \quad (10)$$

Table 10 presents these results. The first two columns present the model including just the Bartik index with and without controls, and the last two columns present the model including the Bartik index interacted with the current LTV dummies. As expected, positive economic shocks in the previous two quarters reduce the likelihood of new foreclosure notices, or neighborhoods with an employment composition that should do worse during the crisis have higher rates of new foreclosure notices. Scaling the estimates by the standard deviation of expected employment growth, 1.6 percentage points, yields a reduction in new filings of approximately 0.15 percentage points relative to a base foreclosure rate of 0.72. Comparing the estimate on the Bartik control to the interaction of the current LTV variables with the Bartik control in columns 3 and 4, the increased likelihood of foreclosure from negative economic shocks effects is concentrated among mortgages with either very high or minimal levels of negative equity. The effects for more typical levels of negative equity between 110 and 150 are closer to zero. Regardless, in all models, the inclusion of controls for tract specific economic shocks has minimal effects on the estimated spillover effects of the stock of mortgage foreclosures with the estimates ranging between 0.086 and 0.088, as compared to between 0.087 and 0.089 in Table 6.

Alternative Neighborhood Geographies

Finally, we examine the robustness of our estimates under alternative geographies looking at spillovers within census block groups and at spillovers within the substantially larger zip codes. Table 11 presents results from regressing likelihood of a first NOT on the block group stock of NOT's using block group predicted share of housing units in negative equity as an instrument. The first stage effects of predicted negative equity fall substantially from about 0.32 to 0.22, potentially due to measurement error given the smaller geography with thinner transaction data,

but the instrument still has substantial power with the F-statistics of 32 and the decline in the IV estimates is relatively modest falling from 0.087 in Table 6 to 0.079. At the zip code level, the results are similarly robust with IV estimates of 0.084, although the power of the instrument is reduced likely due to clustering standard errors at the zip code level. Note that estimate on predicted negative equity at the zip code level increases to 0.57 likely due to the greater volume of transactions resulting in a more accurate prediction of negative equity.

Discussion

This paper provides the some of the first quasi-experimental evidence of broad based neighborhood spillovers from mortgage foreclosures. While existing evidence documents highly localized foreclosure spillovers (often within 250 to 600 feet), these existing studies typically do not examine whether such localized and other broader spillover effects contribute substantially to the high levels of heterogeneity in neighborhood foreclosure rates observed during the crisis. We document substantial spillovers in foreclosure filings at the census tract level. The estimates are robust to both including controls for transaction and mortgage attributes, allowing tracts to differ in the time path of foreclosures during the crisis based on pre-determined tract attributes, and including Bartik style controls for tract level economic shocks based on the industrial employment patterns of tract residents. Further, these effects are broad based arising both in predominantly white, higher income neighborhoods and in neighborhoods with larger numbers of minorities or households in poverty.

Our estimated spillover effects can explain a substantial share of the increase in across neighborhood dispersion in foreclosure rates during the crisis, conservatively 20 percent of the increase for our reduced form estimates and 61 percent of the increase for our IV estimates. Further, the stock of total housing units having experienced foreclosure continues to grow even as

housing prices recover. Given the lag between housing price recovery and declines in the stock of foreclosures, these neighborhood spillover effects likely result in rates of new foreclosure filings that were far more persistent during the recovery than they would have been without the influence of neighborhood spillovers.

REFERENCES

Aizer, A. and J. Currie. 2004. Networks or neighborhoods? Correlations in the use of publicly funded maternity care in California. *Journal of Public Economics* 88(12), 2573-2585.

Anenberg, E. and E. Kung. 2014. Estimates of the size and source of price declines due to nearby foreclosures. *American Economic Review* 104(8), 2527-2551.

Anenberg, E. and E. Kung. 2018. Interest Rates and Housing Market Dynamics in a Housing Search Model. Working Paper.

Baum, C. F., M. E. Schaffer and S. Stillman. 2007. Enhanced Routines for Instrumental Variables/Generalized Method of Moments Estimation and Testing. *The Stata Journal* 7, 465-506.

Bayer, P., F. Ferreira and S. L. Ross. 2016. The Vulnerability of Minority Homeowners in the Housing Boom and Bust. *American Economic Journal: Economic Policy* 8, 1-27.

Bayer, P., F. Ferreira and S. L. Ross. 2017. What Drives Racial and Ethnic Differences in High Cost Mortgages? The Role of High Risk Lenders. *Review of Financial Studies* 31(1), 175-205.

Bayer, P., S. L. Ross and G. Topa. 2008. Place of work and place of residence: Informal hiring networks and labor market outcomes. *Journal of Political Economy* 116, 1150-1196.

Bertrand, M., E. Luttmer and S. Mullainathan. 2000. Network effects and welfare cultures. *The Quarterly Journal of Economics* 115(3), 1019-1055.

Bradley, M. G., A. C. Cutts and W. Liu. 2015. Strategic mortgage default: the effect of neighborhood factors. *Real Estate Economics* 43(2), 271-299.

Brunner, E., S. L. Ross and E. Washington. 2011. Economics and ideology: Causal evidence of the impact of income on support for redistribution and other proposals. *Review of Economics and Statistics* 93, 888-906.

Campbell, J. Y., S. Giglio and P. Pathak. 2011. Forced sales and house prices. *American Economic Review* 101(5), 2108-2131.

Chan, S., M. Gedal, V. Been and A. Haughwout. 2013. The Role of Neighborhood Characteristics in Mortgage Default Risk: Evidence from New York City. *Journal of Housing Economics* 22(2), 100-118.

Chan, S., A. Haughwout and J. Tracy. 2015. How Mortgage Finance Affects the Urban Landscape, In the *Handbook of Urban and Regional Economics Vol. 5A* (Eds. G. Duranton, V. Henderson and W. Strange), 987-1045. Amsterdam: North Holland.

Fletcher, J. M. and S. L. Ross. 2018. Estimating the Effects of Friendship Networks on Health Behaviors of Adolescents. *Health Economics* 27, 1-34.

Fletcher, J. M., S. L Ross and Y. Zhang. 2020. The determinants and consequences of friendship composition. *Journal of Urban Economics*, Article #103241.

Gerardi, K., E. Rosenblatt and P. S. Willen, V. W. Yao. 2015. Foreclosure externalities: some new evidence. *Journal of Urban Economics* 87, 42-56.

Ghent A., R. Hernández-Murillo and M. Owyang. 2014. Differences in subprime loan pricing across races and neighborhoods. *Regional Science and Urban Economics* 48, 199–215.

Glaeser, E. L., B. Sacerdote and J. A. Scheinkman. 1996. Crime and social interactions. *The Quarterly Journal of Economics* 111(2), 507-548.

Goodstein, R. M., P. Hanouna, C. D. Ramirez and C. W. Stahel. 2017. Contagion effects in strategic mortgage defaults. *Journal of Financial Intermediation* 30, 50-60.

Guiso, L., P. Sapienza and L. Zingales. 2013. The determinants of attitudes toward strategic default on mortgages. *The Journal of Finance* 68(4), 1473-1515.

Gupta, A. 2019. Foreclosure contagion and the neighborhood spillover effects of mortgage defaults. *Journal of Finance* 74(5), 2249-2301.

Guren, A.M 2018. House Price Momentum and Strategic Complementarity. *Journal of Political Economy.* 126, 1172–1218.

Guren, A. M. and T. J. McQuade. 2020. How Do Foreclosures Exacerbate Housing Downturns? *Review of Economic Studies* 87(3), 1331-1364.

Harding, J.P., E. Rosenblatt and V.W. Yao. 2009. The contagion effect of foreclosed properties. *Journal of Urban Economics* 66(3), 164-178.

Hartley, D. 2014. The impact of foreclosures on nearby housing prices: supply or dis-amenity? *Regional Science and Urban Economics* 49, 108-117.

Immergluck, D. and G. Smith. 2006. There goes the neighborhood: the effect of single-family mortgage foreclosures on property values. *Housing Policy Debate* 17(1), 57-79.

Ioannides, Y. M. and G. Topa. 2010. Neighborhood effects: Accomplishments and looking beyond them. *Journal of Regional Science* 50(1), 343-362.

Leonard, T. and J. Murdoch. 2009. The neighborhood effects of foreclosure. *Journal of Geographical Systems* 11(4), 317-332.

Lin, Z., E. Rosenblatt and V.W. Yao. 2009. Spillover effects of foreclosures on neighborhood property values. *The Journal of Real Estate Finance and Economics* 38(4), 387-407.

Manski, C.F. 1993. Identification of endogenous social effects: the reflection problem. *Review of Economic Studies* 60, 531-542.

Mian, A., A. Sufi and F. Trebbi. 2015. Foreclosures, House Prices, and the Real Economy. *Journal of Finance* 70, 2587-2634.

Munroe, D. J. and L. Wilse-Samson. 2013. Foreclosure contagion: measurement and mechanisms. Working Paper.

Rogers, W.H., Winter, W. 2009. The impact of foreclosures on neighboring housing sales. *Journal of Real Estate Research* 31(4), 455-479.

Ross, Stephen L. 2011. Social interactions within cities: Neighborhood environments and peer relationships. In *Handbook of Urban Economics and Planning* (Eds. N. Brooks, K. Donaghy and G. Knapp). Oxford, England: Oxford University Press.

Schuetz, J., V. Been, I. Gould. 2008. Neighborhood effects of concentrated mortgage foreclosures, *Journal of Housing Economics* 17(4), 306-319.

Towe, C. and C. Lawley. 2013. The contagion effect of neighboring foreclosures. *American Economic Journal: Economic Policy* 5(2), 313-335.

Table 1. Analysis sample descriptive statistics

Variables	Description	Mean	Standard Deviation
Time Varying Neighborhood Variables			
First NOT	Received first Notice of Trustee Sale	0.0072	0.0847
Fraction NOT	Share pre-crisis sales having received first NOT	0.0421	0.0482
Negative Equity	Share of sales in negative equity	0.3975	0.2490
Pred Neg Equity	Share of sales predicted in negative equity	0.3214	0.2107
Housing Unit Negative Equity and Current Loan to Value Ratio			
Curr LTV 70-90	Current LTV between 0.7 and 0.9	0.2309	0.4214
Curr LTV 90-110	Current LTV between 0.9 and 1.1	0.2281	0.4196
Curr LTV 110-130	Current LTV between 1.1 and 1.3	0.1363	0.3431
Curr LTV 130-150	Current LTV between 1.3 and 1.5	0.0822	0.2747
Curr LTV > 150	Current LTV greater than 1.5	0.0692	0.2537
Transaction and Mortgage Attributees			
Arms length	whether the sale is an arm's length transaction	0.9347	0.2471
Less than 2 yrs old	whether the sale happens within two years after built	0.0774	0.2673
Adjustable Rate	whether the liens are adjustable rate mortgages	0.6361	0.4811
Subordinate debt	whether there are subordinate liens	0.4448	0.4969
Init LTV 80-95	80<Initial Combined LTV<=95	0.2582	0.4376
Init LTV 95-100	95<Initial Combined LTV<=100	0.2174	0.4125
Init LTV 100-110	100<Initial Combined LTV<=110	0.0109	0.1040
Init LTV > 110	Initial Combined LTV>110	0.0129	0.1127
Predetermined Neighborhood Variables (2000 Decennial Census)			
Share Black	Percent individuals who are black	5.2324	6.8493
Share Hispanic	Percent of individuals who are hispanic	20.9302	16.6945
Share Poverty	Percent households in poverty	5.6576	6.0118
Sample Size			
Home Sales	Number of pre-crisis sales	121185.0000	
Sales by Quarters	Number of quarterly NOT observations	1,870,680	

Notes: Means and standard deviations are calculated for the full sample of crisis quarters by pre-crisis period home sales. Home sales are selected as most recent transaction on single-family housing unit between third quarter 2001 and second quarter 2006 that did not have an NOT recorded between the transaction quarter and the second quarter of 2006. The sales by quarter sample is one observations per quarter beginning in the third quarter of 2006 and running until the recording of the first NOT or until the second quarter of 2010 when the data ends for every pre-crisis home sale.

Table 2. Initial loan attributes, by cohort and census tracts partitioned by predicted negative equity increases

Loan Attributes	Census Tracts by Negative Equity Increases		
	Low increase	Middle increase	High increase
Adjustable Rate			
Cohort 02-03	0.384	0.373	0.400
Cohort 03-04	0.636	0.641	0.663
Cohort 04-05	0.786	0.804	0.798
Cohort 05-06	0.728	0.761	0.738
Subordinate Debt			
Cohort 02-03	0.318	0.344	0.335
Cohort 03-04	0.406	0.460	0.442
Cohort 04-05	0.483	0.559	0.519
Cohort 05-06	0.533	0.618	0.570
Unit Less than 2 yrs old			
Cohort 02-03	0.145	0.053	0.133
Cohort 03-04	0.161	0.046	0.127
Cohort 04-05	0.069	0.015	0.036
Cohort 05-06	0.032	0.007	0.016
CLTV > 80			
Cohort 02-03	0.272	0.296	0.288
Cohort 03-04	0.278	0.278	0.267
Cohort 04-05	0.266	0.251	0.238
Cohort 05-06	0.259	0.239	0.213
CLTV > 95			
Cohort 02-03	0.109	0.184	0.193
Cohort 03-04	0.137	0.230	0.240
Cohort 04-05	0.206	0.327	0.296
Cohort 05-06	0.263	0.387	0.341
CLTV > 100			
Cohort 02-03	0.014	0.025	0.029
Cohort 03-04	0.008	0.011	0.014
Cohort 04-05	0.005	0.006	0.009
Cohort 05-06	0.004	0.007	0.008
CLTV > 110			
Cohort 02-03	0.019	0.017	0.018
Cohort 03-04	0.011	0.009	0.011
Cohort 04-05	0.004	0.006	0.007
Cohort 05-06	0.006	0.006	0.005

Notes: The table presents the fraction of transactions in each origination cohort that have the above attribute measures separately for three terciles of census tracts: those tracts having the lowest, middle and highest levels of increase in the predicted share of units in negative equity between the third quarter of 2006 and the second quarter of 2010.

Table 3. Relative Transaction Volume and Average Loan Attributes

Variables	Mortgage/Transaction Attributes	Plus Hedonic Attributes	Plus Hedonic and Housing Price
Mortgage/Transaction Attributes			
Adjustable Rate	-0.079 (0.059)	-0.086 (0.059)	-0.072 (0.060)
Subordinate Debt	0.040 (0.061)	0.032 (0.060)	0.045 (0.059)
Unit Less than 2 yrs old	0.148 (0.186)	0.046 (0.190)	0.051 (0.192)
CLTV > 80	-0.057 (0.063)	-0.038 (0.062)	-0.039 (0.061)
CLTV > 95	-0.052 (0.078)	-0.013 (0.076)	-0.031 (0.076)
CLTV > 100	0.526 (0.555)	0.555 (0.555)	0.465 (0.544)
CLTV > 110	0.0600 (0.165)	0.0516 (0.174)	-0.058 (0.159)
Hedonic Attributes			
Log of Lot Size		0.027 (0.059)	0.023 (0.060)
Log of Square Feet		0.035 (0.105)	0.102 (0.113)
Number of Bedrooms		0.006 (0.041)	0.016 (0.040)
Number of Bathrooms		0.084* (0.048)	0.070 (0.047)
Lot Size Missing		0.501 (0.620)	0.487 (0.609)
Square Feet Missing		1.380 (1.177)	1.307 (1.113)
Market Circumstance			
Housing Price			-0.111 (0.082)
Observations	8,711	8,711	8,675
F-Test [Pr>F]	0.60 [0.75]	0.95 [0.50]	1.05 [0.40]
R-squared	0.464	0.469	0.662

Notes: The Table presents the estimates for a model of the number of mortgage originations in each tract and purchase quarter as a fraction of the median number of quarterly transactions for that tract. The controls are measured as the fraction of mortgages/sales in a tract and purchase quarter that have this attribute. The sample contains all tract by purchase quarter combinations with a positive number of transactions. The model conditions on tract by purchase year and purchase year by quarter fixed effects. The F-tests are intended to detect whether variation in tract mortgage volume over time can be explained by variation in the composition of the mortgages being originated or housing units being sold.

Table 4. Predictiveness of Negative Equity Instrument

	Share in Negative Equity	Fraction NOT
Predicted Negative Equity	0.745***	0.284***
	(0.146)	(0.0682)
Observations	9,376	9,376
R-squared	0.962	0.754

Notes: The table presents estimates of tract share of units in negative equity and fraction of units having received a first NOT on the predicted level of negative equity in a sample of tract by crisis quarter and year. The model contains tract and crisis quarter by year fixed effects.

Table 5. Reduced Form Effects of Negative Equity

Variables	Estimates	
	w/ out controls	w/ controls
Predicted Negative Equity	0.029***	0.029***
	(0.006)	(0.006)
Current LTV 70-90	0.0006***	-0.0007***
	(0.000)	(0.000)
Current LTV 90-110	0.003***	-0.0005*
	(0.000)	(0.000)
Current LTV 110-130	0.007***	0.003***
	(0.001)	(0.001)
Current LTV 130-150	0.012***	0.006***
	(0.001)	(0.001)
Current LTV > 150	0.017***	0.009***
	(0.001)	(0.001)
Observations	1,748,520	1,748,520
R-squared	0.014	0.015

Notes: The table presents estimates of a model for a housing unit
receiving a first NOT in a specific quarter regressed on predicted
fraction of mortgages in the census tract in negative equity and
controls for current LTV. The model controls for tract by purchase
quarter fixed effects and crisis quarter by purchase quarter fixed
effects. The w/ controls column adds the transaction and mortgage
variables from Table 1. Standard errors are clustered at the tract level.

Table 6 Spillover Effects of Foreclosure

Two Stage Least Squares Model of First NOT		
	w/ out controls	w/ controls
Fraction NOT	0.087***	0.089***
	(0.008)	(0.008)

First Stage Model of Fraction NOT		
	w/ out controls	w/ controls
Predicted Negative Equity	0.329***	0.328***
	(0.057)	(0.057)

First Stage F-Stat	32.94	33.27
Observations	1,748,520	1,748,520
First Stage R-squared	0.087	0.094
Second Stage R-squared	0.002	0.003

Notes: The Table presents two stage least squares estimates for whether a pre-crisis home sale has a first NOT in a specific crisis quarter as a function of the fraction of pre-crisis sales receiving a first NOT between the third quarter of 2006 and the current quarter (Fraction NOT). The fraction NOT is instrumented using the predicted share of mortgages in negative equity in each purchase quarter. The model controls for current LTV, census tract by purchase quarter fixed effects and crisis quarter by purchase quarter fixed effects. Panel 1 presents the two stage least squares estimates and panel 2 presents the estimates for the first stage. The w/ controls specification includes the transaction and mortgage attributes from Table 1. Standard errors are clustered at the census tract level.

Table 7 Crisis Year Means and Standard Deviations

Crisis Year	Percent Price Change	First NOT	Fraction NOT	Pred Neg Equity
Year 06-07	-0.026	0.0024	0.0040	0.0568
		(0.0041)	(0.0059)	(0.0145)
Year 07-08	-0.157	0.0076	0.0256	0.2286
		(0.0071)	(0.0236)	(0.0393)
Year 08-09	-0.256	0.0106	0.0669	0.5307
		(0.0110)	(0.0543)	(0.0613)
Year 09-10	0.023	0.0083	0.1033	0.5084
		(0.0062)	(0.0724)	(0.0619)

Notes: Each year represents the 3rd and 4th quarters of the first year listed and the 1st and 2nd quarters of the next year listed. The percent price change uses the average of the quarterly price indices for all four quarters, and the percent change is from the base established in the previous year. The last three columns present means and standard deviations of tract year averages across all transactions from the third quarter of one year to the second quarter of the next. The standard deviation is shown in parentheses.

Table 8 Heterogeneous Effects of Negative Equity and Spillovers

| | Reduced Form Estimates | | | |
	Percent Black	Percent Hispanic	Percent Poverty	All Interactions
Predicted Negative Equity	0.024***	-0.000	0.018***	-0.001
	(0.005)	(0.004)	(0.006)	(0.004)
Pred Neg Equity*Share Black	0.044***			0.019***
	(0.007)			(0.005)
Pred Neg Equity*Share Hispanic		0.044***		0.046***
		(0.003)		(0.004)
Pred Neg Equity*Share Poverty			0.067***	-0.020**
			(0.008)	(0.010)
Observations	1,748,520	1,748,520	1,748,520	1,748,520
R-squared	0.015	0.016	0.015	0.016

| | Two Stage Least Squares Estimates | | | |
	Percent Black	Percent Hispanic	Percent Poverty	All Interactions
Fraction NOT	0.095***	0.085***	0.087***	0.089***
	(0.010)	(0.030)	(0.013)	(0.034)
Fr NOT*Percent Black	-0.056**			-0.054
	(0.028)			(0.044)
Fr NOT*Percent Hispanic		0.008		0.004
		(0.042)		(0.050)
Fr NOT*Percent Poverty			0.015	0.022
			(0.038)	(0.042)
First Stage F-Stat	19.34	4.46	10.91	2.40
Observations	1,748,520	1,748,520	1,748,520	1,748,520
R-squared	0.003	0.003	0.003	0.003

Notes. The Table presents the reduced form and two stage least squares estimates for whether a pre-crisis home sale has a first NOT in a specific crisis quarter as a function of the fraction of pre-crisis sales receiving a first NOT between the third quarter of 2006 and the current quarter (Fraction NOT) and this variable interacted with pre-determined tract attributes. In the two stage least squares estimates, the fraction NOT and its interactions are instrumented using the predicted share of mortgages in negative equity in each purchase quarter and that share interacted with the tract attributes. The model controls for current LTV, census tract by purchase quarter fixed effects and crisis quarter by purchase quarter fixed effects. Panel 1 presents the reduced form estimates and panel 2 presents the estimates for the two stage least squares estimates. Standard errors are clustered at the census tract level.

Table 9 Controlling for Tract Specific Foreclosure Trends

Reduced Form Model of First NOT		
	w/ out controls	w/ controls
Predicted Negative Equity	0.008**	0.008**
	(0.004)	(0.004)
Observations	1,748,520	1,748,520
R-squared	0.015	0.016
Two Stage Least Squares Model of First NOT		
	w/ out controls	w/ controls
Fraction NOT	0.087***	0.079***
	(0.027)	(0.026)
First Stage Model of Fraction NOT		
	w/ out controls	w/ controls
Predicted Negative Equity	0.097***	0.098***
	(0.034)	(0.034)
First Stage F-Stat	8.096	8.273
Observations	1,748,520	1,748,520
First Stage R-squared	0.903	0.903
Second Stage R-squared	0.002	0.003

Notes: The Table presents reduced form two stage least squares estimates for whether a pre-crisis home sale has a first NOT in a specific crisis quarter as a function of the fraction of pre-crisis sales receiving a first NOT between the third quarter of 2006 and the current quarter (Fraction NOT). The fraction NOT is instrumented using the predicted share of mortgages in negative equity in each purchase quarter. The model controls for current LTV, census tract by purchase quarter fixed effects and crisis quarter by purchase quarter fixed effects plus the interaction of pre-determined tract attributes with current year fixed effects. Panel 1 presents the reduced form estimates, panel 2presents the two stage least squares estimates and panel 3 presents the estimates for the first stage. The w/ controls specification includes the transaction and mortgage attributes from Table 1. Standard errors are clustered at the census tract level.

Table 10 Controlling for a Bartik Index of Tract Specific Economic Shocks

Variables	Estimates		Estimates	
	w/ out controls	w/ controls	w/ out controls	w/ controls
Fraction NOT	0.086***	0.088***	0.086***	0.088***
	(0.009)	(0.009)	(0.009)	(0.009)
Bartik Index	-0.092*	-0.107**	-0.182***	-0.110**
	(0.055)	(0.054)	(0.052)	(0.052)
Bartik*Current LTV 70-90			-0.014	-0.075***
			(0.010)	(0.010)
Bartik*Current LTV 90-110			0.013	-0.091***
			(0.014)	(0.014)
Bartik*Current LTV 110-130			0.153***	0.058***
			(0.021)	(0.019)
Bartik*Current LTV 130-150			0.294***	0.187***
			(0.026)	(0.025)
Bartik*Current LTV > 150			0.002	-0.055*
			(0.032)	(0.030)
First Stage F-Stat	28.499	28.789	29.386	29.438
Observations	1,748,520	1,748,520	1,748,520	1,748,520
R-squared	0.002	0.003	0.002	0.003

Notes: The Table presents two stage least squares estimates for whether a pre-crisis home sale has a first NOT in a specific crisis quarter as a function of the fraction of pre-crisis sales receiving a first NOT between the third quarter of 2006 and the current quarter (Fraction NOT). The fraction NOT is instrumented using the predicted share of mortgages in negative equity in each purchase quarter. The model controls for current LTV, census tract by purchase quarter fixed effects and crisis quarter by purchase quarter fixed effects. The first two columns present estimates from a model that includes the Bartik Index without and with controls for mortgage attributes, and the second two columns also include interactions between the current LTV dummy variables and the Bartik Index. Standard errors are clustered at the census tract level.

Table 11 Spillover Effects of Foreclosure (Block group)

Reduced Form Model of First NOT		
	w/ out controls	w/ controls
Predicted Negative Equity	0.0173***	0.0174***
	(0.00466)	(0.00483)
Observations	1,748,520	1,748,520
R-squared	0.025	0.025
Two Stage Least Squares Model of First NOT		
	w/ out controls	w/ controls
Fraction NOT	0.0791***	0.0799***
	(0.0115)	(0.0118)
First Stage Model of Fraction NOT		
	w/ out controls	w/ controls
Predicted Negative Equity	0.218***	0.218***
	(0.0384)	(0.0380)
First Stage F-Stat	32.226	32.833
Observations	1,748,520	1,748,520
First Stage R-squared	0.058	0.065
Second Stage R-squared	0.003	0.004

Notes: The Table presents reduced form two stage least squares estimates for whether a pre-crisis home sale has a first NOT in a specific crisis quarter as a function of the fraction of pre-crisis sales receiving a first NOT between the third quarter of 2006 and the current quarter (Fraction NOT). The fraction NOT is instrumented using the predicted share of mortgages in negative equity in each purchase quarter. The model controls for current LTV, block group by purchase quarter fixed effects and crisis quarter by purchase quarter fixed effects plus the interaction of pre-determined block group attributes with current year fixed effects. Panel 1 presents the reduced form estimates, panel 2presents the two stage least squares estimates and panel 3 presents the estimates for the first stage. The w/ controls specification includes the transaction and mortgage attributes from Table 1. Standard errors are clustered at the block group level.

Table 12 Spillover Effects of Foreclosure (Zip code)

Reduced Form Model of First NOT		
	w/ out controls	w/ controls
Predicted Negative Equity	0.0481***	0.0497***
	(0.0102)	(0.0111)
Observations	1,748,520	1,748,520
R-squared	0.009	0.010
Two Stage Least Squares Model of First NOT		
	w/ out controls	w/ controls
Fraction NOT	0.0839***	0.0869***
	(0.0104)	(0.00860)
First Stage Model of Fraction NOT		
	w/ out controls	w/ controls
Predicted Negative Equity	0.573***	0.571***
	(0.132)	(0.130)
First Stage F-Stat	18.890	19.218
Observations	1,748,520	1,748,520
First Stage R-squared	0.143	0.149
Second Stage R-squared	0.001	0.002

Notes: The Table presents reduced form two stage least squares estimates for whether a pre-crisis home sale has a first NOT in a specific crisis quarter as a function of the fraction of pre-crisis sales receiving a first NOT between the third quarter of 2006 and the current quarter (Fraction NOT). The fraction NOT is instrumented using the predicted share of mortgages in negative equity in each purchase quarter. The model controls for current LTV, zip code by purchase quarter fixed effects and crisis quarter by purchase quarter fixed effects plus the interaction of pre-determined zip code attributes with current year fixed effects. Panel 1 presents the reduced form estimates, panel 2presents the two stage least squares estimates and panel 3 presents the estimates for the first stage. The w/ controls specification includes the transaction and mortgage attributes from Table 1. Standard errors are clustered at the zip code level.